

(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11)

EP 1 362 910 A1

(12)

**EUROPEAN PATENT APPLICATION**  
published in accordance with Art. 158(3) EPC

(43) Date of publication:  
19.11.2003 Bulletin 2003/47

(51) Int Cl.7: C12N 15/09, C12P 21/02,  
C12Q 1/02, G01N 33/48

(21) Application number: 02700643.6

(86) International application number:  
PCT/JP02/01503

(22) Date of filing: 20.02.2002

(87) International publication number:  
WO 02/066634 (29.08.2002 Gazette 2002/35)

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE TR  
Designated Extension States:  
AL LT LV MK RO SI

(30) Priority: 20.02.2001 JP 2001044221

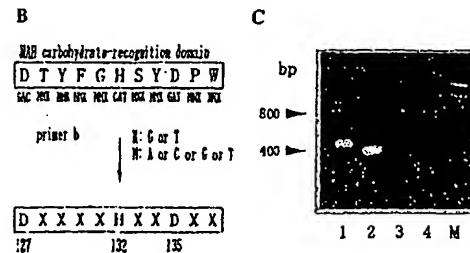
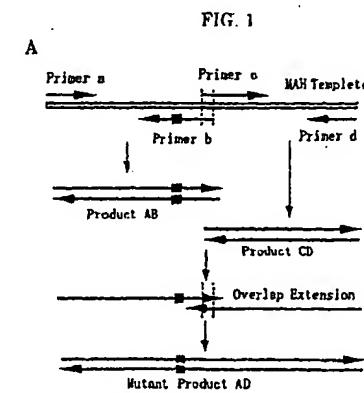
(71) Applicants:  
• Irimura, Tatsuro  
Tokyo 154-0015 (JP)  
• Matsumoto, Mariko  
Kawasaki-shi, Kanagawa 211-0954 (JP)

(72) Inventors:  
• Irimura, Tatsuro  
Setagaya-ku, Tokyo 154-0015 (JP)  
• Matsumoto, Mariko  
Kawasaki-shi, Kanagawa 211-0954 (JP)  
• Yim, Mijung  
Seoul 151-818 (KR)  
• Ono, Takashi  
Nakakoma-gun, Yamanashi 400-0316 (JP)

(74) Representative: Baldock, Sharon Claire et al  
BOULT WADE TENNANT,  
Verulam Gardens  
70 Gray's Inn Road  
London WC1X 8BT (GB)

**(54) LECTINS FOR ANALYZING SUGAR CHAINS AND METHOD OF USING THE SAME**

(57) It is expected that lectins of a needed number of types appropriately corresponding the variation of sugar chains, if available, are highly useful as tools for analyzing various sugar chains. From a lectin library containing plural types of lectins or an arbitrary mass of lectins containing plural types of lectins, lectins needed for analyzing sugar chains are selected to thereby construct a lectin sub-library. Using the lectin sub-library thus constructed, cells having sugar chains are identified. Moreover, an indication method whereby cells can be conveniently identified is provided.



**Description****Technical Field of the Invention**

5 [0001] The present invention relates to a technique of analyzing cell surface sugar chains or the like by using lectins. In particular, the present invention relates to a sugar chains analyzing library and sub-library comprising lectins that can analyze sugar chains, and relates to a method of analyzing sugar chains by using the library and sub-library for analyzing sugar chains. Further, the present invention relates to a library and sub-library of lectins for use in identifying cells, which can easily and accurately identify the type of particular cells. The present invention also relates to a method 10 of identifying the type of cells by using the library and sub-library for identifying cells.

**Background Art**

15 [0002] In recent years, it has been said that if the code recorded in sugar chains could be decoded, new horizons of biology would be opened. As a tool for analyzing various sugar chains or as a molecular candidate decoding such a code in the living body, carbohydrates-binding proteins (lectins) can be mentioned. However, each individual lectin cannot give macroscopic information for appropriately reflecting the variety of sugar chains.

20 [0003] That is, it can duly be expected that if a number of different types of lectins could be obtained, sufficient enough to adequately correspond to the variation of sugar chains, they can constitute a useful tool for analyzing various types of sugar chains sugar chains.

**Disclosure of Invention**

25 [0004] In view of the problem described above, a lectin sub-library is constructed in the present invention by selecting lectins necessary for analyzing sugar chains, from a lectin library comprising plural types of lectins. By using the lectin sub-library thus constructed, cells having sugar chains are identified. Further, a method of display capable of easily identifying cells is provided herein. That is, sugar chains are hardly specified by using one type of lectin, but they can be identified, by using a lectin sub-library comprising a combination of different types of lectins. Further, even in the case that the sugar chains structure cannot be fully specified in details, the lectin sub-library can still be used as a tool 30 to identify and to make a comparison of sugar chains wherein a difference in cells (including a difference in type and a difference between normality and abnormality) represented by a difference of sugar chains, or a difference in the glycosylation modification of proteins can be recognized, or a substance carrying sugar chains as markers can be discriminated. On the other hand, when the relationship is clear between a lectin sub-library and cells, glycosylation modification of glycoproteins, or sugar chains markers, the difference in cells, the difference of glycosylation modification of glycoproteins, sugar chains markers can be identified from the relationship, without elucidating the actual structure 35 of sugar chains.

[0005] In particular, the present invention provides:

40 (1) A lectin sub-library for analyzing sugar chains, comprising a series of lectins each having a modification at a predetermined site on the lectin, or plural types of lectins obtained by a panning with cells or pseudo-cells having predetermined sugar chains. The panning may be conducted for a lectin library composed of plural types of lectins. Incidentally, insofar as the function of a sub-library can still be maintained as a whole, some elements (lectins) of the sub-library may be omitted. Similarly, insofar as the function of a library can still be maintained as a whole, some elements of the library may also be omitted.

45 [0006] The present invention further provides:

50 (2) A method of analyzing sugar chains, which comprises using a cell-identifying lectin sub-library comprising a series of lectins each having a modification in a predetermined site in the lectin, or plural types of lectins obtained by pannings with cells or pseudo-cells having sugar chains.

(3) A tool for diagnosing diseases, which comprises using the lectin sub-library as described in the above-mentioned (1).

55 (4) A solid carrier wherein lectin-immobilizing regions each of which contains a lectin selected to identify sugar chains, are arranged in a suitable order for identifying sugar chains. The lectin-immobilizing region may be a well or a simple spot, e.g. a region not particularly delimitated by an external structure, or may be a part or place where lectins can be immobilized. The solid carrier may be in a form of chip or plate or any other forms of solid material capable of forming the lectin-immobilizing regions thereon.

(5) The solid carrier according to the above-mentioned (4), characterized in that the selected lectin has been

obtained by gene mutation.

(6) The solid carrier according to the above-mentioned (5), characterized in that the predetermined lectin is MAH lectin.

5 (7) A method of display to identify cell surface sugar chains binding pattern by using a solid carrier wherein lectin-immobilizing regions are arranged in a suitable order for identifying/analysing sugar chains, and each region contains a lectin selected for identifying/analysing sugar chains.

(8) The method of display according to the above-mentioned (7), characterized in that the selected lectin has been obtained by gene mutation.

(9) The method of display according to the above-mentioned (8), wherein the predetermined lectin is MAH lectin.

10 (10) A solid carrier for analyzing/comparing sugar chains, wherein different types of lectins having specific sugar chains binding specificities to the cell surface sugar chains are immobilized in a predetermined order.

15 (11) A sugar chains-comparing tool and method comprising using at least one lectin selected from a plurality of lectins known to bind to at least 2 types of sugar chains. The selected lectins show a greater difference with respect to the binding activity to the 2 types of carbohydrates. When a value "A" that represents the ability of a certain lectin to bind to one type of sugar chain, is greater than a value "B" that represents the ability of the lectin to bind to another type of sugar chain, the relative difference in the binding ability can be expressed as  $(A - B)/(A + B)/2$ . The greater difference with respect to the binding capability may refer to those relative differences greater than the average relative difference in the binding ability of all the lectins, or when the relative differences are arranged in the order of from the greatest to smallest differences, the great relative difference may refer to those relative differences of from the greatest one to a predetermined one (for example, 1, 2, 3, 4 etc.). The sugar chains-comparing tool may contain a solid carrier having lectins arranged in a suitable order to compare sugar chains.

20 (12) The sugar chains-comparing tool and method according to the above-mentioned (11), further comprising using at least one lectin showing a smaller difference with respect to the ability in binding to one type of sugar chains relative to another type of sugar chains in said 2 types of sugar chains. The small relative difference may refer to those relative differences smaller than the average relative difference in the binding ability of all the lectins, or when the relative differences are arranged in the order of from the smallest to greatest differences, the small relative difference may refer to those relative differences of from the smallest one to a predetermined one (for example, 1, 2, 3, 4 etc.).

30 [0007] Generally, the term "lectin" is a generic name of proteins specifically recognizing and binding to sugar chains, which are roughly divided into animal lectins and plant lectins. Among the plant lectins, leguminous lectins constitute a large lectin family. The lectin is composed of a dimer or tetramer of subunits each having a molecular weight of 30,000. Each subunit has one sugar chains-recognizing site. Lectin is sometime defined in a narrow sense as "a multivalent, carbohydrates binding protein excluding enzymes or antibodies, or defined as a glycoprotein", or can be defined as "a carbohydrates recognizing protein binding and cross-linking to carbohydrates" in a broad sense, and has thus the ability to identify and analyze many types of a carbohydrates substantially. Particularly, the variation of carbohydrates may reflect the variation of cells having such carbohydrates on the surface. Therefore, cells can be identified by identifying/analyzing their cell surface carbohydrates.

35 [0008] Unlike antibodies, lectins are found in every organ and tissue in almost all living things. Further, the structures of lectins are varied, while the structures of antibodies are similar.

40 [0009] In the present invention, leguminous lectins are chosen as one example of a large number of various lectins to explain the present invention as completed. However, lectins other than leguminous lectins can be used in a similar manner, and as a matter of course, the present invention is not limited to such leguminous lectins.

45 [0010] For example, lectins having specific binding properties to certain sugar chains may be selected and used without any modification as a lectin library, or mutated lectins at least a part of which carbohydrates-recognizing regions are genetically engineered by random mutation can consist a lectin library as well.

50 Specifically, the terms "certain sugar chains" encompass the cases where sugar chains are defined at least partially, or sugar chains specified by a process of synthesizing/engineering the carbohydrates without knowing the exact structure. Further, the diversity (or an abnormality) in glycosylation of glycoproteins may result different sugar chains structure and this may also be an example of the subjects having "certain sugar chains" and glycoprotein itself can be used to select specific lectins from a lectin library. Further, cell surface sugar chains are characteristic and sugar chains are different from cells to cells, and thus "certain sugar chains" may be read as "certain cells" and "certain cells" can be used to select specific lectins from a lectin library. Further, if a cell has different cell surface sugar chains depending on the environment at the time of glycosylation, the cell cycle(for example a stage before differentiation or a differentiated state) may be identified. This is because glycosylation patterns depends on the cell cycle, environment and cell surface carbohydrates depends on diseases, aging etc.

55 [0011] The phrase "lectins having specific binding properties" is used because not all the mutated lectins have the same binding specificities, in terms of the binding preference to specific carbohydrates. When a parental lectin has a

specific binding specificity to a specific carbohydrate structure, there is a high possibility that the mutated lectins generated from the gene mutation of its carbohydrates-binding site have a similar carbohydrate binding specificities pattern. However, mutated lectins have different binding preference each other and also different from that of the parental lectin. In this way, a plurality of lectins would be obtained, each showing different preference and capability to bind to a certain sugar chain.

[0012] The phrase "at least a part of which carbohydrates-recognizing regions" is given because lectins recognize carbohydrates by cross-linking to its carbohydrate-binding site, a certain region in its whole structure. The terms "at least a part" is given because gene mutation corresponding to the whole structure of the lectin is not always necessary.

[0013] The terms "genetically engineered" is given because chemical modification of an intact lectin is simply difficult and inefficient. Therefore; in the method, genes encoding "carbohydrate recognition regions" are genetically mutated and is transfected into a suitable host such as *Escherichia coli*, to produce a mutated lectin. The term "random mutation" means random genetic mutation, and also predetermined gene mutations to produce random lectins. The mutation used herein means the gene mutation by swapping nucleic acid base without changing the length of the gene or by making the length of the gene longer or shorter. This also applies to the mutated lectins produced from the mutations.

[0014] A lectin collection containing a plurality of lectins obtained in this manner may be called as a lectin library. A plurality of lectins is contained therein because a single lectin is not always sufficient for identifying various sugar chains.

[0015] Site-directed mutagenesis is performed because only the carbohydrate-binding site in the lectin molecule engages in the recognition of carbohydrates. By modifying such a specific site, a preferable lectin library would effectively be obtained. Of course, there may be the case where the effect of the lectin library can be evaluated from some experimental results, and thus it is not always necessary for the mutated site to be limited to such a specific site.

[0016] The method of identifying sugar chains (or a cell having a certain sugar chains) by a lectin library involves examining the binding of specific sugar chains to the lectins consisted the lectin library as described above, and then using the same lectin library to classify the carbohydrates to be identified, by evaluation the binding to the lectin library with two ranks, i.e. the presence of a binding ability (showing bonding) or the absence of a binding ability (showing no bonding) or with three ranks, i.e. the two ranks plus an intermediate rank or by multi-stage evaluation (or analog evaluation) with increased ranks, thereby distinguishing the type of the sugar chains in terms of the ability to bind to a plurality of lectins. Further, the method may also involve the identification/analysis of the binding ability by analyzing its binding pattern.

[0017] There may be various methods of judging the binding ability, and for example, the binding ability may be judged by comparison with a standard sample previously prepared.

[0018] As used herein, the "cells or pseudo-cells having predetermined sugar chains" have sugar chains, or mimic of the cell surface sugar chains. Accordingly, when a lectin sub-library is prepared for cell having carbohydrate A, carbohydrate A or mimic-A can usually be used to get an appropriate sub-library. The terms "analyzing sugar chain" may include identification of cells having sugar chain on the surface, diagnosis of sugar chain-modified serum glycoproteins (glycoform), diagnosis of diseases, and other analysis, and the "method of analyzing sugar chains" includes methods for conducting the above. The tool for diagnosing diseases by using a lectin sub-library for analyzing sugar chain, comprising a series of lectins prepared by introducing modifications at a predetermined site on the lectin, or plural types of lectins selected by panning with cells or pseudo-cells having predetermined sugar chains refers to a tool (diagnostic reagent, diagnostic apparatus, etc.) to judge the state of a disease (or health) when sugar chains on cell surface or glycoproteins are different in the two states, and a tool to examine the changes of the sugar chains through the lectin-binding specificities corresponding to the changes. It is not always necessary to determine the exact structure (or state) of the sugar chain.

[0019] The phrase "lectin-immobilizing regions are arranged in a suitable order for identifying/analyzing sugar chains, and the region contains a lectin selected for identifying/analyzing sugar chains" has a broad concept encompassing "wells or spots containing lectins selected for identifying cells, are arranged in a suitable and appropriate order for the cell identification", and the phrase "wells or spots containing lectins selected for identifying cells are arranged in a suitable and appropriate order for the cell identification" involves arranging wells or spots containing the selected lectins in predetermined positions so as to make their visual pattern easily recognizable for the identification, and if a standard pattern has been established, the phrase involves arranging the wells or spots in positions corresponding to those in the standard pattern. As used herein, "display" may be attained by either direct or indirect display. The chip can be used to provide a visual pattern obtained by sugar chain structure information, and on the basis of this information, the cells can be identified. To identify cells by enormous pattern information, visual judgment by the naked eye or any other methods is often insufficient, and thus the cells may be identified by cluster analysis with a computer. In the cluster analysis, a computer may be used to examine which part indicates a significant difference between subjects to be compared (for example, a standard and an object of examination or a control and an object of comparison), and general cluster analysis software can also be used.

[0020] In the phrase "solid carrier for analyzing/comparing sugar chains wherein different types of lectins are immobilized in a predetermined order", the terms "immobilized in a predetermined order" means that the position at which

different types of lectin is immobilized, may be arranged depending on intention of using sugar chain information for cell analysis or disease diagnosis. This means, the order to immobilize lectins may be changed depending on the cell types and diseases. Similarly, in the "chip wherein wells containing lectins selected for identifying cells are arranged in a suitable and appropriate order for the cell identification", the order may be suitably varied depending on the intention of using sugar chain information for cell analysis or disease diagnosis. The order may preferably be optimized so as to be adapted to the intended cells or diseases. The solid carrier may be in a form of chip or a plate, or any other form of solid materials that can provide lectin-immobilizing regions. That is, the "solid carrier for analyzing/comapring sugar chains" includes "a test plate for analyzing sugar chains". The "test plate for analyzing sugar chains" includes a test paper using a paper such as a filter paper, or a glass plate or a substrate made of other material having predetermined lectins immobilized thereon. Other material for the substrate may include synthetic resin (including plastics), metals (including platinum, silver, copper, gold, silicon, etc.), mica and mixtures thereof. The "test plate for analyzing sugar chains" can also be used as a diagnostic reagent.

#### Brief Description of Drawings

##### [0021]

Fig. 1 shows the introduction of mutation in MAH carbohydrate recognition domain.

Fig. 2 is a photograph showing expression of recombinant MAH mutant lectins in transformed *Escherichia coli* induced by isopropyl thiogalactoside.

Fig. 3 is a list showing the amino acid sequences of the carbohydrate recognition domains of 16 mutant MAH clones.

Fig. 4 shows the hemagglutination activity of 16 mutant lectins (mutant MAH lectins) on erythrocytes from different types of animals.

Fig. 5 shows a mutation site in the carbohydrate binding site of MAH.

Fig. 6 is a graph showing the ability of wild-type MAH-expressing phage to bind to anti-MAH antibody.

Fig. 7 is a graph showing the ability of wild-type MAH-expressing phage to bind to human glycophorin.

Fig. 8 is a graph showing the sugar chain specificity of human erythrocyte-specific clones obtained by panning.

Fig. 9 shows the results of binding activity of the clones (1 to 18) to Caco-2 cells measured by a conventional cell-ELISA method.

Fig. 10 shows the results of binding activity of the clones (19 to 35) to Caco-2 cells measured by a conventional cell-ELISA method.

Fig. 11 is a graph showing binding patterns of differentiated or undifferentiated Caco-2 cells to the clones in a lectin library.

Fig. 12 are graphs showing binding patterns of various cells to commercially available lectins.

Fig. 13 illustrates one example of lectin chips and a method of using the same (for use in IgA discrimination).

Fig. 14 illustrates one example of lectin chips and a method of using the same (for use in profiling of differentiation of osteoblast).

#### Detailed Explanation of the Invention

[0022] Hereinafter, the present invention is explained in more detail by reference to working examples.

[0023] At least a part of a region considered to correspond to the carbohydrates-binding site of *Maackia amurensis hemagglutinin* (MAH), which is a leguminous lectin specific for sialic acid-containing carbohydrates, is mutated at random by genetic engineering techniques, and from the resulting randomly mutated lectins, an artificial library was prepared capable of identifying the sugar chains variations in different type of sugar chains. Further, by using the resulting artificial lectin library, it was attempted to identify different type of cells based on the binding pattern to the lectins in the library. Then, from the artificial lectin library, a screening system established for selecting the lectins specific for biologically important sugar chains.

[0024] It is known that a vast variety of complex carbohydrates having slightly different structures, exist among mammalian cells. Formation of lectins capable of distinguishing such various structures is very useful.

[0025] Such lectins can be obtained from *Maackia amurensis hemagglutinin* (MAH) by genetic engineering techniques.

[0026] MAH can recognize a carbohydrates whose hydrocarbon sequence comprises a sialic acid residue, e.g. Neu5Aco2-3Galβ1-3(Neu5Aco2-6)GalNAc (4). Another isolectin engineering *Maackia amurensis leukoagglutinin* (MAL) can specifically recognize a sequence Neu5Aco2-3Galβ1-4GlcNAc (5), and both lectins are unique among other lectins including leguminous lectins.

[0027] MAH is a dimer of subunits having a relative molecular mass of 29,000.

[0028] A nucleotide sequence of cDNA encoding MAH and an amino acid sequence deduced therefrom show that

MAH consists of 287 amino acids and has a single peptide consisting of 30 amino acids.

[0029] The putative carbohydrate-recognizing domain of MAH was identified by comparing its amino acid sequence with amino acid sequences of other leguminous lectins (7) and by a study of genetic mutation on this domain defined by amino acids endowing MAH with binding characteristics.

5 [0030] Further, those observations were confirmed from a computer model of a three-dimensional structure of MAH containing Neu5Aco2-3Gal $\beta$ 1-3 (Neu5Aco2-6) GalNAc (8).

[0031] Figs. 1A to C are drawings showing the introduction of a mutation in MAH carbohydrate-recognition domain. A is an illustration showing a partial overlap extension method for introducing a mutation. 1B is a drawing showing the carbohydrates-recognition domain of MAH, where mutation is introduced at X sites. 1C is a photograph of agarose gel 10 electrophoresis showing identification of a DNA of mutant MAH. In 1C, lane 1 is a product AB (up to 400 bp), lane 2 is a product CD (up to 400 bp), lane 3 is a product AD (up to 800 pb), lane 4 is a DNA of wild-type MAH (up to 800 pb), M is DNA marker (100 bp DNA Ladder, Life Technologies, Inc., MD, USA).

15 [0032] Fig. 2 is a photograph showing expression of recombinant MAH mutant lectins in transformed *Escherichia coli* induced by isopropyl thiogalactoside. 2A is a photograph showing the result of polyacrylamide gel electrophoresis analysis of recombinant mutant MAH lectins. 2B is a photograph showing the result of Western blotting analysis of recombinant mutant MAH lectins with an anti-MAH polyclonal antibody. In Fig. 2, lane N represents *E. coli* carrying pGEX-2T plasmid, lane P represents *E. coli* carrying pGEX-2T/wild-type MAH plasmid, and lanes 1 to 16 represent *E. coli* carrying pGEX-2T/mutant MAH plasmids (corresponding to clones 1 to 16, respectively). Further, their relative molecular masses are shown in the left, and the arrow indicates recombinant MAH mutant lectins.

20 [0033] Fig. 3 is a list showing the result of deduced amino sequences of sugar-recognition domains in 16 mutant MAH clones.

[0034] Fig. 4 are graphs showing the result of the hemagglutination activities of 16 mutant lectins on erythrocytes from different kinds of animals. The activity of each mutant lectin is shown in terms of minimum hemagglutination concentration relative to maximum hemagglutination concentration assumed to be 1. Nos. 1 to 16 refer to clones 1 to 25 16, and No. 17 refers to a negative control (GST).

<Example 1>

[Preparation of an artificial lectin library]

30 [0035] On the basis of information obtained from the amino acid sequences of various leguminous lectins and carbohydrate-binding sites thereof, a nucleotide sequence encoding the carbohydrate-binding site at the 127- to 137-positions in 285 amino acids of MAH lectin was randomized by overlap extension PCR using synthetic oligonucleotides primers. However, aspartic acid 135 considered to be essential for interaction with sialic acid and aspartic acid 127 35 and histidine 132 considered to be essential for coordination with a metal ion were fixed (Fig. 5). On determining the DNA sequence by the dye terminator method, it was confirmed that the obtained mutant MAHs, excepting the initially fixed 3 amino acids, have different amino acid sequences with each other, indicating that the randomization of MAH lectin has been attained.

40 [Discrimination and identification of cells by the lectin library]

[0036] The prepared lectin library was integrated in a vector pGEX and expressed as GST fusion protein in *Escherichia coli*. Fifty clones were selected at random, and from these clones, 16 clones considered to express the protein were obtained. Western blotting of the proteins with an anti-MAH polyclonal antibody indicated that all the clones remained reactive. After the determination of the DNA sequences of these clones, it was confirmed that the amino acid 45 sequences of the sugar-binding sites in all clones, excepting the fixed amino acids, had been changed.

[0037] By utilizing a difference in sugar-binding specificity among the resulting 16 clones, it was attempted to distinguish and identify erythrocytes of different kinds of animals, on which sugar chains comprising a sialic acid but having different structure is considered to exist. Fig. 4 shows the result of the experiments where the agglutination titer of each 50 lectin of the 16 clones is compared. In the graphs, the data were standardized by taking the maximum agglutination titer shown by the most reactive lectin for the respective erythrocytes as "1" (full-scale). The respective erythrocytes were found to show inherent characteristics in terms of the relative agglutination titer against the 16 mutant lectins.

[0038] This experiment suggests that the difference in the type of cell can be distinguished by referring to the binding 55 patterns with respect to a plurality of lectins.

[Preparation of a phage library]

[0039] The artificial lectin library would be further developed by selecting lectins specific for a biologically useful

carbohydrate chain. For this selection, a system capable of expression cloning and dealing with a large number of libraries should be used. In this context, a phage display system was introduced. It was attempted to use an expression vector  $\lambda$ foo and then a T7 phage display system. However, the growth of lectin-expressing phage was delayed as compared with the counterpart phage not carrying the gene, and it was found difficult to concentrate the desired phage by panning. (However, this does not necessarily mean that the expression vector  $\lambda$ foo or the T7 phage display system is excluded from the present invention.)

[0040] Incidentally, panning refers to a technique of recovering phage by utilizing the ability of a protein expressed on the phage to bind to the binding partner (for example, the binding ability between antigen and antibody, lectin and carbohydrate, receptor and ligand, etc.) For example, a substance that binds to the protein expressed on phage is immobilized on a plate, then a phage solution is added to the plate thereby allowing the expressed proteins, phage and other materials to bind to the said substance, then the phage solution is discharged, and the expressed proteins, phage and other materials thus binding to the substance are recovered. The recovered phage can be used to infect with *Escherichia coli* to proliferate and then subjected again to (the above-described) panning to recover highly binding phage.

[0041] Then, it was attempted to use phageimids pComb3 and pComb8. It was first examined whether a lectin expressed on the surface of both the phages had an activity or not by expressing wild-type MAH. When the phages considered to express wild-type MAH were examined for their reactivity with an anti-MAH antibody, both phages were bound specifically to the anti-MAH antibody. pComb8 expressing a plurality of MAH proteins showed a stronger binding ability than pComb3 expressing a single MAH protein (Fig. 6).

[0042] It was examined whether the wild-type MAH expressed on the phages was maintaining the carbohydrate-binding activity by measuring the ability to agglutinate erythrocytes. Both the phages considered to express wild-type MAH agglutinated untreated erythrocytes, but did not agglutinate erythrocytes treated with sialidase. The minimum titer of both the phages necessary for hemagglutination was about  $1 \times 10^{10}$  cfu. Thus, it was revealed that the activity of MAH can be maintained on the phage, and therefore MAH having a carbohydrate binding activity can be expressed on the pComb phage.

[0043] The activities of both phages were further examined in more detail in terms of the ability to bind to human glycophorin (sialic acid-containing glycoprotein). It was unexpectedly found that pComb3 expressing a single MAH protein per phage had a higher activity (Fig. 7). The minimum titer of pComb3 phage detectable by ELISA in terms of the ability to bind to human glycophorin was about  $1 \times 10^7$  cfu.

[Selection of specific lectins from a phage library]

[0044] On the basis of the above results, the pComb3 phageimid was used to prepare a lectin library having mutated primary structures. For preparation of the library in a feasible range, 3 amino acids that were considered to be essential for interaction with sialic acid and for coordination with a metal ion, were fixed. Other two amino acids that are highly conserved among other leguminous lectins, were also fixed. The remaining 6 amino acids were mutated at random in the amino acid sequence of the MAH sugar-binding site. As a result, a library of about  $1 \times 10^9$  lectins considered to cover the theoretically estimated number of lectins of  $6.25 \times 10^7$  could be prepared.

[0045] For establishing a panning system for the lectin library of pComb3 phage, a panning with human erythrocytes that strongly bind to the wild-type MAH was performed. The recovery in the third panning was about 10 times as high as that in the first panning, thus revealing that the phage having an ability to bind to human erythrocytes was concentrated. From the phages obtained after the third panning, several clones that specifically agglutinate human erythrocytes but shows different carbohydrates specificity were obtained (Fig. 8). In this way, a screening system could be established for selecting lectins specific for biologically important carbohydrates from a library of lectins having mutated primary structures.

[Mutagenesis of the sugar chain recognition site of MAH]

[0046] The carbohydrates recognition site of MAH was randomized by site-directed mutagenesis (9, 10) utilizing overlap extension procedure. In the overlap extension, the following 4 primers were used.

Table 1. Four primers

Primer a : 5'-ccccggatccacatcagatgagctttct-3' (bp 89-105)

Primer b : 5'-atgtcgataattggatctnnmnnatcmnnmnnatgmnnmn  
nmnnmnmngtcaaactctacagc-3' (bp457-519)

(N represents A, T, G or C, and M represents A or C)

Primer c :: 5'-gatccaaattatcgacatatcgaaattgat-3' (bp 502-531)

Primer d 5'-cccgaaattcagatcatgcagtgtaaacg-3' (bp 847-531)

[0047] MAH cDNA fragment was PCR amplified by using primer pairs of a and b or c and d with AmpliTaq Gold DNA polymerase (Perkin Elmer). Standard PCR conditions were used. Specifically, the reaction mixture was incubated at 95°C for 9 minutes and at 94°C for 1 minute and then subjected to 30 cycles each consisting of incubation at 94°C for 1 minute, at 54°C for 1 minute and at 72°C for 1 minute followed by one cycle of incubation at 72°C for 10 minutes, whereby all amplified fragments were extended certainly to the 3' end. The 3'-end of the PCR product was polished with a plus DNA polymerase (Stratagene La Jolla, California). The PCR product was purified and quantified by preparative agarose gel electrophoresis. Equimolar amounts of the two PCR fragments were bound to each other by 7 cycles each consisting of incubation at 94°C for 1 minute and at 63°C for 4 minutes. The product was used as template DNA in second reaction using primers a and d, whereby a full-length mutant MAH cDNA was prepared.

### [Expression and purification of GST fusion protein]

**[0048]** The PCR product containing the full-length mutant MAH cDNA was digested with BamHI and EcoRI and then ligated to pGEX-2T previously digested with BamHI and EcoRI. *E. coli* BL21 strain containing the expression plasmid was cultured in a 2x yeast-tryptone medium supplemented with 50 µg/ml ampicillin, and grown at 37°C until the absorbance at 660 nm reached 0.8. Expression of GST fusion protein was induced by adding isopropyl β-D-thiogalactoside (final concentration: 1 mM) and then the cells were incubated at 37°C for 3 to 4 hours. A pellet of the microbial cells was suspended in Tris buffered physiological saline (TBS), and sonicated. The insoluble fraction was washed twice with TBS and dissolved in 8 M urea in TBS. Thus solubilized inclusion body was refolded by 50-fold dilution with TBS/1 mM MnCl<sub>2</sub>/1 mM CaCl<sub>2</sub>. The supernatant was applied onto a glutathione-Sepharose 4B column (Amersham Pharmacia Biotech UK Ltd., England), and the GST fusion protein was purified according to the manufacturer's instructions.

#### [Detection of GST protein by Western blotting]

[0049] The microbial lysate (0.5 µg protein/ml) was separated by sodium dodecylsulfate-polyacrylamide gel electrophoresis, transferred onto a nitrocellulose membrane, screened by polyclonal anti-MAH antibody and then incubated with an alkali phosphatase-bound anti-rabbit IgG antibody (ZYMED, San Francisco, California). According to manufacturer's instructions, the bound antibody was visualized with an alkali phosphatase substrate kit SK-5200 (Vector Laboratories, Beringam (phonetic), California).

## [Determination of DNA sequences of recombinant MAH mutants]

**[0050]** DNA sequences of 16 clones selected at random were determined by using a sense primer 5' -ttcttgaccac-ctgattctc-3' and an antisense primer 5'-ccggcgtaatccaaatccat-3' with ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, California, US).

### [Preparation of erythrocytes]

[0051] Human venous blood of the O type was taken from a blood donor by using a previously heparin-treated hypodermic syringe. The heparin-treated blood was transferred into a glass cylinder and left until the erythrocytes were precipitated by gravity. After leukocyte-rich serum and buffy coat were removed, the erythrocyte layer was washed 3 times by centrifugation with phosphate buffered saline (140 mM NaCl, 2.7 mM KCl, 10 mM  $\text{Na}_2\text{HPO}_4$ , 1.8 mM  $\text{KHPO}_4$  (pH 7.3)), and the outermost layer was removed carefully each time of washing after centrifugation. It was confirmed that the erythrocytes obtained in this manner did not contain leukocytes, as will be described below.

[0052] A recombinant protein purified from each of the 16 mutant lectins was diluted twice and put to a U-bottomed plate (Nunc, NY, USA), and to each dilution (0.25ml) was added an equal volume of 2% erythrocytes. The

mixture was left for 1 hour at room temperature and then examined for hemagglutination.

[0053] The following gives more specific disclosure.

[Mutagenesis of the sugar chain recognition site of MAH]

5

[0054] The overlap extension method (9, 10) was used to introduce a mutation into the sugar chain recognition site of MAH. The scheme (strategy) of this method is shown in Fig. 1.

10

[0055] Using MAH cDNA cloned in pGEX-2T as a template for PCR amplification, two different DNA fragments (AB and CD) were formed. The fragment AB was amplified with primers a and b. Because Asp 135, Asp 127 and His 132

15

were considered to participate in binding of MAH to sialic acid, primer b was designed such that while Asp 127, His 132 and Asp 135 residues were stored, the sequence of the sugar chain recognition site of MAH was mutated (Fig. 1 B).

[0056] The fragment CD was amplified with primers c and d. The 5' end of the primer c had a 12-bp sequence found in the 3' end of the fragment AB. The fragments AB and CD can be ligated to each other via this common 12-bp sequence by the overlap extension method to form an AD product harboring the mutant MAH cDNA. The length (up to 800 bp) of the mutant MAH cDNA was confirmed by agarose gel electrophoresis (Fig. 1C).

[Expression of MAH mutant lectins and elucidation of characteristics thereof]

20

[0057] The mutant MAH cDNA was inserted into a BamHI-EcoRI site in *E. coli* expression vector pGEX-2T containing a glutathione-S-transferase (GST) gene upstream of the insertion site. GST/mutant lectin fusion proteins were expressed by using *Escherichia coli* BL21 as a host strain. The ability of 50 randomly selected clones to express the recombinant mutant lectins was examined by polyacrylamide gel electrophoresis. GST/mutant lectin fusion proteins from 16 clones were electrophoresed to a position corresponding to an apparent molecular weight of 55,600, which is predicted with an molecular weight of the GST/mutant lectin fusion product (Fig. 1A).

25

[0058] Plasmid DNAs from the 16 clones were sequenced by a dye deoxy chain terminator method (11). Fig. 3 shows the elucidated amino acid sequences of the sugar chain recognition sites of these clones. As expected, the sugar chain recognition sites of the 16 clones were different in amino acid sequence excepting Asp 127, His 132 and Asp 135 residues. Regardless of this difference, all the 16 mutant lectins were detected by Western blotting with an antibody specific for MAH (Fig. 1B).

30

[0059] The 16 mutant lectin fusion proteins were purified by affinity chromatography on a glutathione-Sepharose column. The concentration was determined by a BCA protein assay kit with bovine serum albumin as the standard. Profiles of the purified fusion proteins after polyacrylamide gel electrophoresis in the presence of SDS suggest that all fusion proteins have a molecular weight of 55,600. These proteins were used in hemagglutination assays.

35

[Hemagglutination activity of the mutant lectins]

40

[0060] Hemagglutination assays using human, bovine, equine, porcine and chicken erythrocytes revealed that there was a significant difference in the sugar chain specificity among the 16 mutant lectins and the wild-type MAH (Fig. 1). Sugar chains on the surfaces of these erythrocytes are different in the structure, amount and expression pattern as is evidently shown by comparing the respective recognition patterns by the mutant lectins (Fig. 4). Each type of erythrocyte can be easily identified by using these lectins.

45

[0061] In a radial graph in Fig. 4 similar to a form of bicycle spokes, the 16 different lectins and blank are expressed as straight lines of the same length, and these straight lines are arranged radially from one point at equal angles through 360°. In Fig. 4, one radial graph is prepared for erythrocytes from one type of animal, and this graph shows the discrimination (or characteristics) of the erythrocytes by the respective lectins in a visual manner. That is, the maximum agglutination titer observed is expressed in full scale "1" and plotted in the outermost side in this graph. The agglutination titers of the other lectins are plotted in a similar manner, respectively, and adjacent plots are linked via a straight line. By expressing plots in this manner, the positions of the plots is more easily recognizable than by simple plots, and the resulting graph in a web form permits discrimination (or characteristics) of the erythrocytes by the respective lectins to be easily and visually recognized. The 16 lectins can be arranged so that the similar ones are placed as adjacent as possible, contrastive ones are placed apart at 180°, and so on. It is preferable that the arrangement of the plot of lectins is suitably coordinated depending on the sugar chain to be analyzed. The graph in Fig. 4 shows all the 16 lectins, but the graph may be modified to show 15 or less lectins, wherein some lectins not necessary for discrimination are omitted, or it may of course include 17 or more lectins adding thereto some additional lectins. Further selection of such lectins can be carried out, for example, from the viewpoint of their ability to bind to intended sugar chains. Referring to Fig. 4, clone 15 can be selected for distinguishing bovine erythrocytes from human erythrocytes. This is because clone 15 has an ability to bind strongly to human erythrocytes but is poor in an ability to bind to bovine erythrocytes and the relative difference therebetween is great. Further, clone 13 may also be selected. This is because clone 13 has the

same ability to bind to human erythrocytes and bovine erythrocytes and the relative difference therebetween is small. By selecting lectins showing great and small relative differences respectively for intended erythrocytes, unknown samples of erythrocytes could be judged easily and reliably.

5 [0062] The life cycle of cells in multi-cellular living things is regulated by stimuli from extracellular environments, and such stimuli are mediated at least partially via sugar chain and its recognizing molecules. It has been considered that, as in the case of receptors specific for antigens in an adapted immune system, cell surface carbohydrates have a wider variety. However, there is a limit to the variation of natural lectins capable of recognizing such tremendous variety of carbohydrates/oligosaccharides.

10 [0063] In the present invention, a method is provided that enables the preparation of a number of novel lectins (novel lectin library) having various specificities. The 16 mutant MAH lectins were produced. Although their characteristics in terms of carbohydrate binding specificity has not fully been revealed, it is noteworthy that the mutants can agglutinate human and animal erythrocytes with highly specific patterns. The specificity of lectins to distinguish different animal erythrocytes from one another has not been achieved except by using specific antibody thereagainst. The present invention is to prove that a panel of genetically engineered lectins having various specificities is useful for distinguishing 15 the type of different cells.

20 [0064] In the last several years, DNA arrays techniques were established as a new technology for distinguishing the type of difference cells thorough profiling of gene expression (15). However, the phenotype of cells cannot fully be determined only on the basis of gene expression. Posttranslational modification by adding sugar chains and control of protein localization are considered to have essential roles in determining the actual stage in the differentiation of 25 certain type of cells. Accordingly, different cells should be distinguished by describing the state of glucan on the surfaces of the cells. Use of a panel of lectins having various specificities is the most convenient and effective means for achievement thereof. Development of a novel lectin library can significantly expand the applicability of lectins to ultimately make the identification of cells accurate and reliable. Although the expression of carbohydrates are very low, and their structure is complicated, the lectin library and/or lectin sub-library according to the present invention can be used to provide a convenient discrimination (or identification) method that can also be used in clinical fields.

<Example 2>

[Cell discrimination and identification targeted at Caco-2 cells]

30 [0065] A mutant MAH lectin library consisting of loop D mutated MAH lectins was subjected to panning with differentiated Caco-2 cells as standard cells, and clones used as probes for cell discrimination and identification of the target Caco-2 cells were recovered. After the second panning, the concentration of the binding clones was considered insufficient, and thus in the third panning, the cells were fixed with glutaraldehyde, and the reaction time in the panning 35 procedure was changed into o/n.

[Determination of DNA sequences of recovered clones]

40 [0066] After the third panning, 64 clones were selected at random and analyzed for their sequence. The 64 clones consist of 35 clones having the intended mutations, 7 clones having a stop codon, 17 clones undergoing frame-shift, and 5 clones that could not be analyzed. Among 35 clones having the intended mutations, there were no clones having the same sequence as that of the wild type and the clones have no common sequence at the mutated site (Table 2).

45

50

55

5

10

15

20

25

30

35

40

45

50

## 3rd clone analysis

Table 2. Sequences of Clones

wild-type	Asp	Thr	Tyr	Pho	Gly	His	Ser	Tyr	Asp	Pro	Trp	wild-type	Asp	Thr	Pho	Gly	His	Ser	Tyr	Asp	Pro	Trp		
	GAC	act	taC	TTC	GGG	cat	AGT	TAT	GAT	CCC	TGG		GAC	act	TAC	TTC	GGG	cat	AGT	TAT	GAT	CCC	TGG	
clone 1												clone 19												
	gac	act	taC	CAG	GCG	cat	ATT	ACG	cat	CAT	TGG		gac	act	TAC	ATG	TGG	cat	QAT	GTG	cat	CGT	TCT	
				Leu	Pro	Ser	Leu								Leu									
clone 2				gac	act	taC	TTC	CCT	cat	TCG	TTC		clone 20											
														gac	act	TAC	GCT	ATG	cat	TGT	TGG	cat	CGT	ATG
clone 3				gac	act	taC	QAG	CGT	cat	ACT	GGG													
													clone 21											
clone 4				gac	act	taC	TCC	GCT	cat	CCT	GCT													
													clone 22											
clone 5				gac	act	taC	GCT	CTT	cat	CCG	GCG													
													clone 23											
clone 6				gac	act	taC	TCT	CCT	cat	TCT	TGG													
													clone 24											
clone 7				gac	act	taC	CTT	TGT	cat	TCT	TGG													
													clone 25											
clone 8				gac	act	taC	GGG	GGG	cat	TCG	CCT													
													clone 26											
clone 9				gac	act	taC	TCQ	GCG	cat	TCT	TGG													
													clone 27											
clone 10				gac	act	taC	AGG	GTT	cat	GCG	TCT													
													clone 28											
clone 11				gac	act	taC	CTQ	GTQ	cat	QTT	CGG													
													clone 29											
clone 12				gac	act	taC	CTT	GGG	cat	TAT	CCT													
													clone 30											
clone 13				gac	act	taC	CAQ	CCT	cat	QCA	CTT													
													clone 31											
clone 14				gac	act	taC	CQA	QCT	cat	TTO	QGO													
													clone 32											
clone 15				gac	act	taC	QCG	QAG	cat	QCT	TAT													
													clone 33											
clone 16				gac	act	taC	CGT	TAT	cat	QCT	ACT													
													clone 34											
clone 17				gac	act	taC	TGT	CCT	cat	TCT	CTG													
													clone 35											

55 A large number of clones different from the wild type has been obtained, these 35 clones can be used as probes for discrimination and identification of cells. On the other hand, since no common sequence was detected among these clones, and clones having a stop codon or undergoing frame-shift were recovered, the concentrating efficiency in the panning procedure is considered as not so high in this experiment.

[Measurement of the activity of the recovered-clones]

[0067] Using a lysate of each clone, the activity thereof to bind to Caco-2 cells was measured by a conventional cell-ELISA method (Figs. 9 and 10). All of the 35 clones had an activity to bind to both differentiated and undifferentiated Caco-2 cells. When binding patterns of the respective clones were plotted with respect to the differentiated and undifferentiated Caco-2 cells, a binding pattern appeared to be unique to Caco-2 cells was obtained and there was a slight difference in the pattern between the differentiated and undifferentiated cells (Fig. 11). These results indicate that a difference in the pattern between the differentiated and undifferentiated cells can be observed even in the same kind of cells. This type of graph in Fig. 11 is superior in judging a slight difference because the results of both the differentiated and undifferentiated cells can be collectively shown in one graph.

<Example 3>

[0068] Profiling of various cells was conducted as well. The commercially available purified lectins (PNA, ABA, VVA-B4, SNA, UEA-I, WGA, PHA-L4) were immobilized directly onto wells and sample cells were added to the wells to detect the binding thereof (reverse cell-ELISA method) (Fig. 12). As can be seen from the graphs in Fig. 12, a binding pattern unique to each kind of cell is obtained and expressed in a graph in a radial form similar to a form of bicycle spokes, whereby characteristics of the cell can be grasped. The manner of plotting the graph in Fig. 12 is the same as that in Figs. 4 and 11 described above.

[Discrimination of IgA glycoforms]

[0069] An application of the lectin library to a method of distinguishing IgA glycoforms is described below.

1) A genetically mutated lectin library is prepared by mutating an amino acid sequence of the sugar chain recognition site of MAH lectin (Maackia amurensis hemagglutinin lectin).  
 2) Lectins having higher affinity for an IgA from a patient having IgA nephropathy are selected by the panning method, to prepare a lectin sub-library.  
 3) Lectins from the lectin sublibrary obtained in step 2) above, are immobilized on a microtiter plate. Optionally, lectins sufficiently reflecting a difference between the IgA from the patients having IgA nephropathy and IgA from healthy persons may be specifically selected from the lectin sub-library obtained in step 2) above, to prepare the plate.  
 4) Discrimination by a lectin library of IgA glycoforms: Binding patterns, on the lectin plate, of serum IgA from the patients having IgA nephropathy and serum IgA from healthy persons are compared and analyzed.  
 5) Investigation of assay conditions for serum analysis: Sugar chains are added enzymatically or chemically to IgA from healthy persons, to prepare artificial IgA. The artificial IgA is mixed with serum from healthy persons and subjected to pattern analysis by using the lectin plate. The influence of other serum proteins present in serum may be examined to optimize assay conditions.

[0070] IgA nephropathy is a disease in which IgA is precipitated in renal glomerulus, and the IgA molecule is different in the sugar chain structure on its hinge region between the IgAs from patients and healthy persons. The analysis of the sugar chain structure is conventionally carried out by purifying serum IgA, separating the hinge-containing peptide fragment by a treatment with trypsin, and measuring its mass by a mass spectrometer. The hinge region refers to a region between two H-chain constant domains constituting an immunoglobulin molecule, and it is assumed that in patients with IgA nephropathy, there is an abnormality in an O-linked sugar chain on the hinge region in IgA. Further, an O-linked sugar chain having, as a fundamental skeleton, a structure composed of N-acetyl galactosamine (GalNAc), galactose (Gal) and outermost sialic acid, is also referred to as a mucin type sugar chain. The O-linked type sugar chain can significantly influence on such as intercellular interactions, cellular infiltration, adhesion etc.

[0071] The detailed mechanism of the method is as follows: The O-linked type sugar chains can be attached onto serine and threonine residues at 5 distinctive sites in the amino acid sequence of the hinge region in IgA, where patients with IgA nephropathy have an abnormality in the O-linked sugar chain. There are 6 different patterns of the O-linked sugar chains, and thus  $6^5$  (i.e. 7,776) different sugar chain structures can theoretically occur in relation to the attaching positions in the hinge region. It is difficult to examine every possibility by a complicated sugar chain analysis conducted conventionally. However, the pattern analysis would be feasible by observing ON/OFF signals toward lectins that recognize the O-linked sugar chain. By examining ON/OFF toward one lectin, two sugar chain linkage modes can be determined, and with "n" lectins given,  $2^n$  patterns would be recognizable. Accordingly, the number of lectins necessary for analyzing 7,776 sugar chain abnormalities in the hinge is theoretically 13. This theoretical number is elucidated based on an assumption that there are lectins capable of clearly reflecting the structural and positional information of

the sugar chain. However, if information from several hundred lectins is available and the lectins can be blotted on a substrate, there is a high possibility that the structural and positional information of the sugar chain in the IgA hinge can be obtained. To obtain a useful lectin library, a group of lectins that clearly reflects the difference in between serum IgAs from patients and from healthy persons, can be identified by a cluster analysis to construct an effective lectin sub-library. Using the lectin library, a difference in glycoform between serum IgA from patients with IgA nephropathy and serum IgA from healthy persons is analyzed by the pattern information (Fig. 13). Fig. 13 shows that the profiling of IgAl, a glycoprotein containing a plurality of O-glucans, can easily be carried out. There is applicability not only to analysis of IgA from patients with severe diseases such as those in chronic renal insufficiency but also to early diagnosis of potential patients with a slight symptom not diagnosed as IgA nephropathy.

10 [Method of distinguishing osteoblast subgroup]

15 [0072] One example of the application of the lectin library to a method of distinguishing osteoblasts or to a method of detecting a subgroup different in the stage of differentiation is described below.

- 1) A genetically mutated lectin library is prepared by mutating an amino acid sequence of the sugar recognition site of MAH lectin (*Maackia amurensis* hemagglutinin lectin).
- 2) Lectins having high affinity for the prepared osteoblasts are selected from the lectin library by panning to prepare a lectin sub-library.
- 20 3) Lectins from the lectin sub-library obtained in the step 2) above, are immobilized on a microtiter plate. Optionally, lectins sufficiently reflecting the stages of differentiation are selected from the lectin sub-library obtained in the step 2) above, to prepare the lectin plate.
- 4) Introduction of the differentiation of osteoblasts and discrimination by the lectin library: Mesenchymal stem cells are cultured and separated. Specifically, the mesenchyme stem cells are cultured and then differentiated into osteoblasts, and the cells on the 5th day, 10th day, 15th day and 20th day after the initiation of differentiation are separated.
- 25 5) The cells separated at each stage in the differentiation process are analyzed with the lectin plate. The correlation between the sugar chain structure on the surface of the cell and the ability to form bone is examined. The ability to form bone is determined by measuring the activity of bone-type alkali phosphatase and simultaneously measuring the content of osteocalcin (preparation of a standard). That is, the separated cells in the form of a complex with a b-TCP block serving as an anchor for the osteoblasts are transplanted subcutaneously in the back of a rat, and 4 weeks and 8 weeks after transplantation, the cells are removed and examined for (i) osteocalcin content and (ii) bone-type alkali phosphatase activity.
- 30 6) Cells whose differentiation process is not identified/unrevealed are analyzed with the lectin plate and compared with the standard pattern (see Fig. 14). Fig. 14 shows that profiling of differentiation of osteoblasts can be easily conducted.

35 [0073] As used herein, the mesenchymal stem cells are stem cells which among those differentiated into tissues or organs, occur in the bone marrow, and it is confirmed that the mesenchymal stem cells are differentiated into cells in bone, cartilage, fat, heart, nerves, liver etc., and have a potential similar to that of embryonic stem cells (ES cells) which can be differentiated into almost all tissues. The mesenchymal stem cells can be used as an anchor for culturing an artificial material based on calcium phosphate called b-TCP (b-tricalcium phosphate) excellent in affinity, absorptivity, and bone conductivity in the living body.

40 [0074] According to the present invention, the lectin library was found to be useful for discrimination and identification of cells.

45 [0075] By analyzing cell surface sugar chain by this new method, identification of cells, which has solely been depended on gene expression only, or antibodies whose antigen recognition is limited, be easily carried out, and thus this method would be useful for development of cell transplantation and cell therapy. Further, this study succeeded in expressing an artificial lectin library on the surface of phageimid-type phage. Because the method of selecting lectins having different specificities from a mutant lectin library was established, it is expected that this system can be used for obtaining lectins having novel sugar chain specificity absent in existing lectins or monoclonal antibodies.

50 [0076] Further, the lectin library of the present invention can be used to provide a tool for detection of IgA glycoforms or for discrimination/identification of an osteoblast subgroup. That is, the lectin library can provide not only a diagnostic reagent for easily and rapidly analyzing glycoforms of various serum proteins including IgA, but also a standard design tool for guaranteeing qualities of cells necessary for bringing regenerative medical treatment or cellular medical treatment into a practical stage. For example, these tools can be applied to early discovery of diseases, to accurate understanding of a morbid state and to therapeutic/prophylactic medicines, for example by analyzing glycoforms of immunoglobulins in rheumatism and autoimmune diseases, glycoforms of specific hormones in patients with cancers (that

is, change in sugar chains of chorionic gonadotropin in ovarian cancer) and glycoforms of proteins (change in sugar chains of alpha-fetoprotein at a stage of from hepatitis to hepatoma).

[0077] Because an antibody is characterized by being produced depending on a difference in species, an antibody specific for fibroblasts is hardly produced. Therefore, fibroblasts derived from bone marrow do not have a clear so-called cell surface marker, and there is no method of directly assaying osteoblasts. Even at present time, bioassays with animals or indirect bioassays reflecting the activity of osteoblasts are employed. In such a field, a lectin chip using the lectin library of the present invention would be feasible in the discrimination/identification of osteoblasts, particularly in the discrimination/identification of an osteoblast subgroup at different stages. By the same approach, it can duly be expected that the lectin library of the invention can be applied to the analysis of the dendritic cells derived from bone marrow and the vascular endothelial cells derived from bone marrow, and therefore can be used as a standard tool for securing qualities of cells used in regenerative medical treatment.

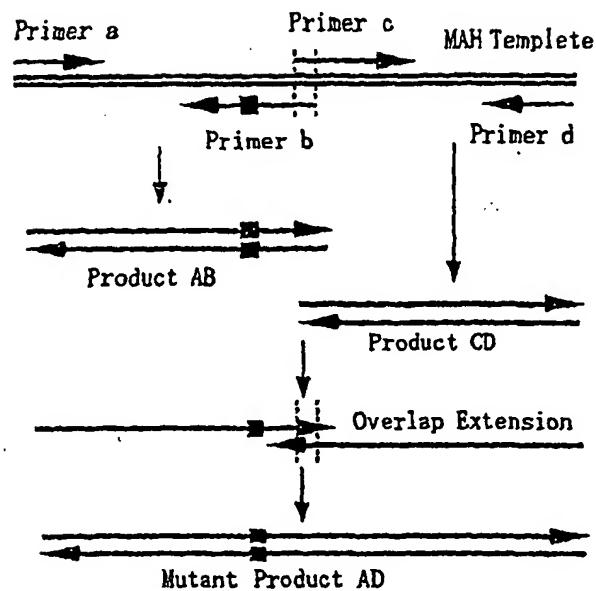
[0078] A network for remote diagnosis of cancers at present is based on pharmacological diagnosis, but it is examined to add a genetic information thereto. On finding the considerable involvement of the O-linked sugar chain in the inter-cellular interactions, cell infiltration, adhesion etc., it would be possible to achieve more detailed remote diagnosis of cancers.

### Claims

1. A lectin sub-library for analyzing sugar chains, comprising a series of lectins each having a modification at a pre-determined site on the lectin, or plural types of lectins selected by a panning with cells or pseudo-cells having predetermined sugar chains.
2. A method of analyzing sugar chains, which comprises using the lectin sub-library according to claim 1.
3. A tool for diagnosing diseases, comprising the lectin sub-library according to claim 1.
4. A solid carrier for analyzing sugar chains, wherein different types of lectins selected for distinguishing sugar chains, from the lectin sub-library according to claim 1, are immobilized in a predetermined order.
5. A solid carrier wherein lectin-immobilizing regions are arranged in a suitable order for distinguishing sugar chains, and the region contains a lectin selected for analyzing sugar chains.
6. The solid carrier according to claim 5, **characterized in that** the selected lectin is obtained by mutating a part of the gene encoding a lectin.
7. The solid carrier according to claim 6, **characterized in that** the lectin to be mutated is MAH lectin.
8. A method of displaying a cell discrimination pattern by using a solid carrier wherein lectin-immobilizing regions are arranged in a suitable order for analyzing sugar chains, and the region contains a lectin selected for distinguishing sugar chains.
9. The method according to claim 8, **characterized in that** the selected lectin is obtained by mutating a part of the gene encoding a lectin.
10. The method according to claim 9, **characterized in that** the lectin to be mutated is MAH lectin.
11. A sugar chain-comparing tool and method comprising using at least one lectin selected from a plurality of lectins known to bind to at least 2 types of sugar chains, and the selected lectin shows a greater difference with respect to the ability in binding to one type of sugar chain relative to another type of sugar chain in said 2 types of sugar chain.
12. A sugar chain-comparing tool and method according to claim 11, further comprising using at least one lectin showing a smaller difference with respect to the ability in binding to one type of sugar chain relative to another type of sugar chain in said 2 types of sugar chain.

FIG. 1

A



B

## MAH carbohydrate-recognition domain

D	T	Y	F	G	H	S	Y	D	P	W
GAC	NNK	NNK	NNK	NNK	CAT	NNK	NNK	CAT	NNK	NNK

primer b

↓  
 K: G or T  
 N: A or C or G or T

D	X	X	X	X	H	X	X	D	X	X
127					132			135		

C

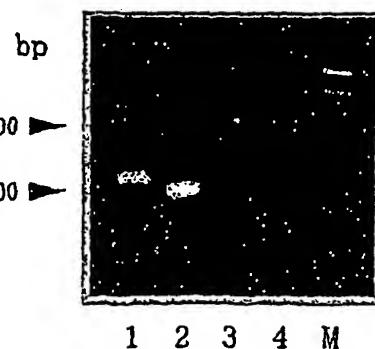


FIG. 2

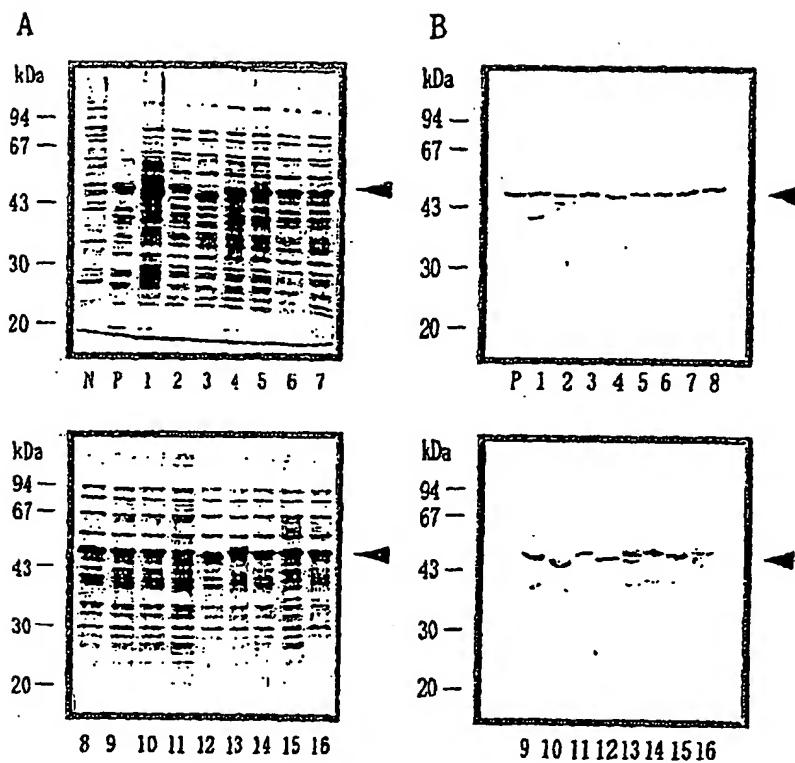


FIG. 3

	130										135			
MAH : D	T	Y	F	G	H	S	Y	D	P	W				
Clone1 : D	N	L	I	Y	H	I	N	D	M	A				
Clone2 : D	R	P	L	T	H	S	D	D	P	V				
Clone3 : D	A	N	V	L	H	K	L	D	R	R				
Clone4 : D	V	P	N	V	H	K	I	D	Y	R				
Clone5 : D	E	S	H	C	H	V	E	D	V	E				
Clone6 : D	V	M	P	L	H	L	Q	D	T	T				
Clone7 : D	T	S	L	P	H	V	V	D	V	G				
Clone8 : D	N	L	T	G	H	L	P	D	S	G				
Clone9 : D	M	H	V	V	H	N	S	D	N	L				
Clone10: D	L	H	S	N	H	T	L	D	K	Q				
Clone11: D	P	K	P	L	H	A	I	D	S	S				
Clone12: D	V	S	H	W	H	T	Q	D	Y	R				
Clone13: D	T	E	L	L	H	T	N	D	H	K				
Clone14: D	S	A	S	N	H	V	A	D	E	R				
Clone15: D	S	S	G	I	H	D	L	D	G	R				
Clone16: D	P	P	M	A	H	Y	R	D	W	P				

FIG. 4

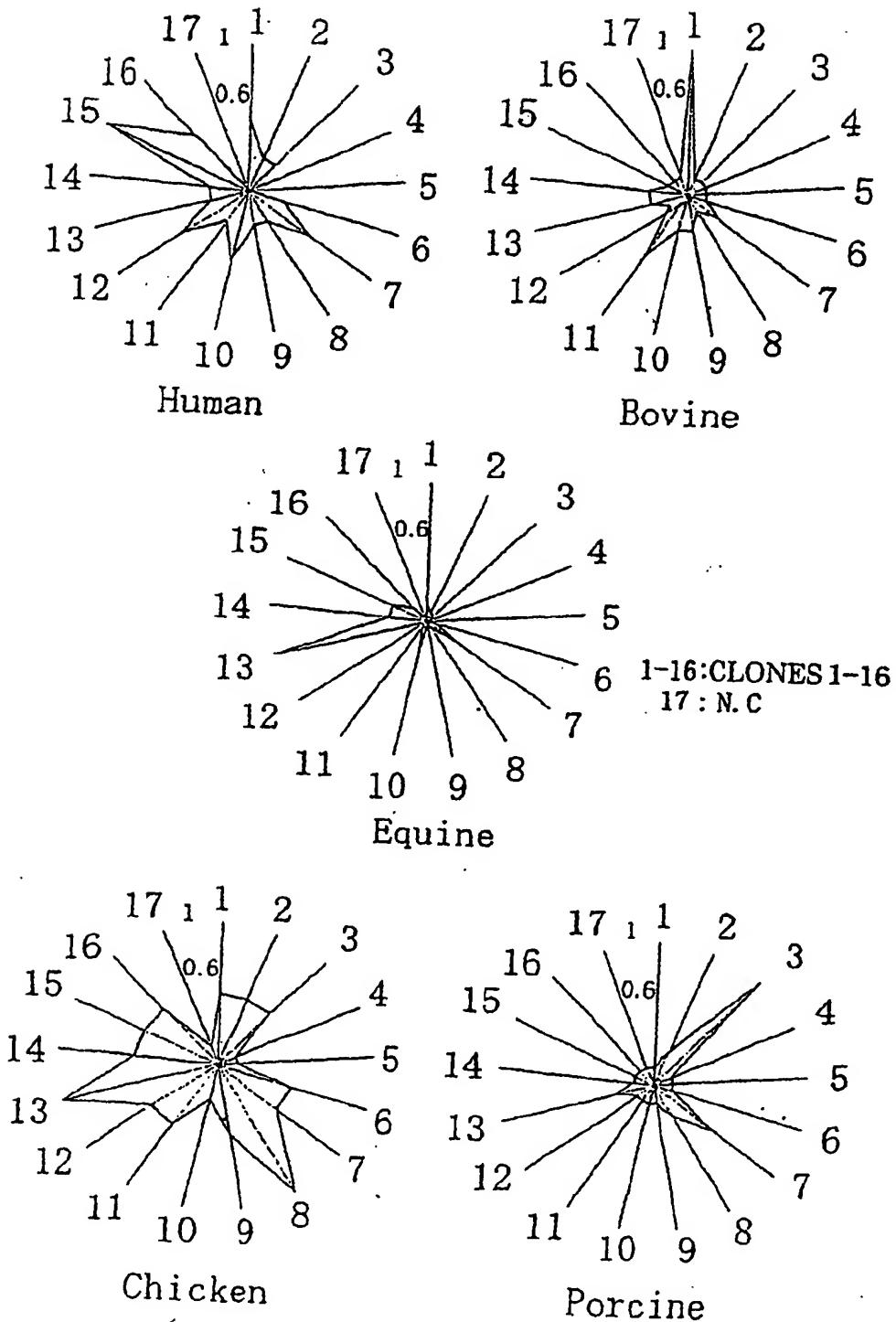


FIG. 5

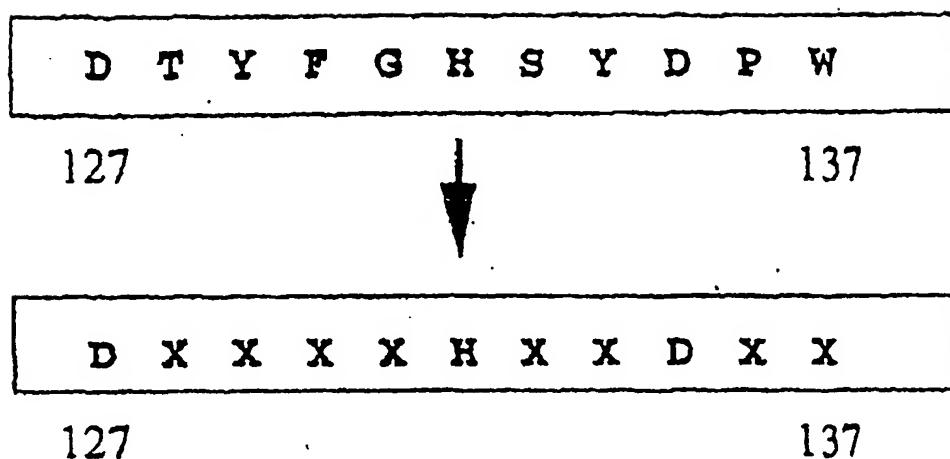


FIG.6

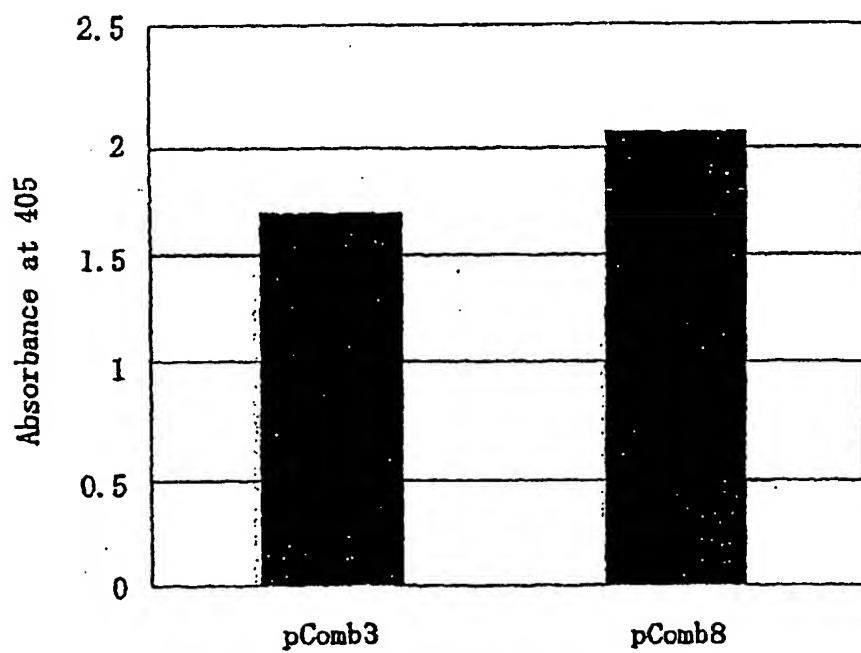


FIG. 7

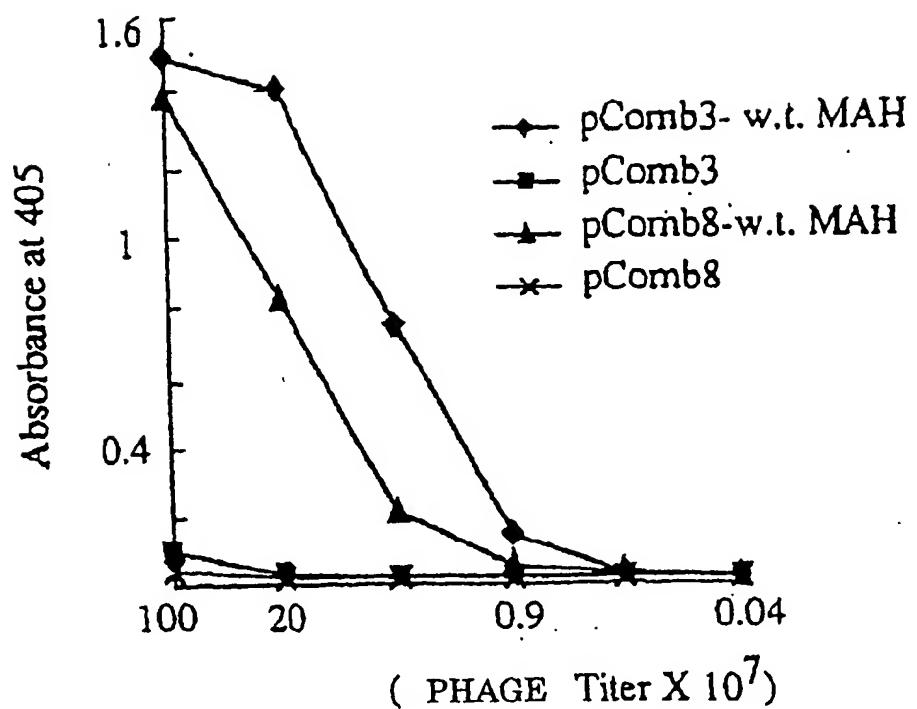
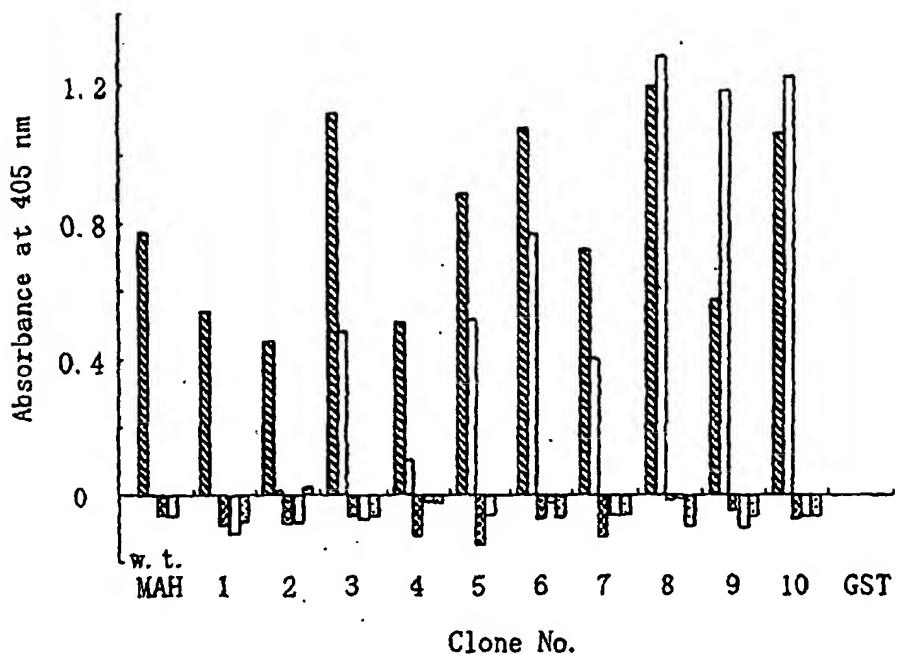


FIG. 8



*hatched bars, glycophorin; gray bars, Neu5Aca2-3Galβ1-3GalNAca-sp-biotin; checked bars, GalNAcal-3GalNAcβ-sp-biotin; white bars, Neu5Aca2-6GalNAca-sp-biotin; dotted bars, 3' sialyllactosamine-BSA.*

FIG. 9

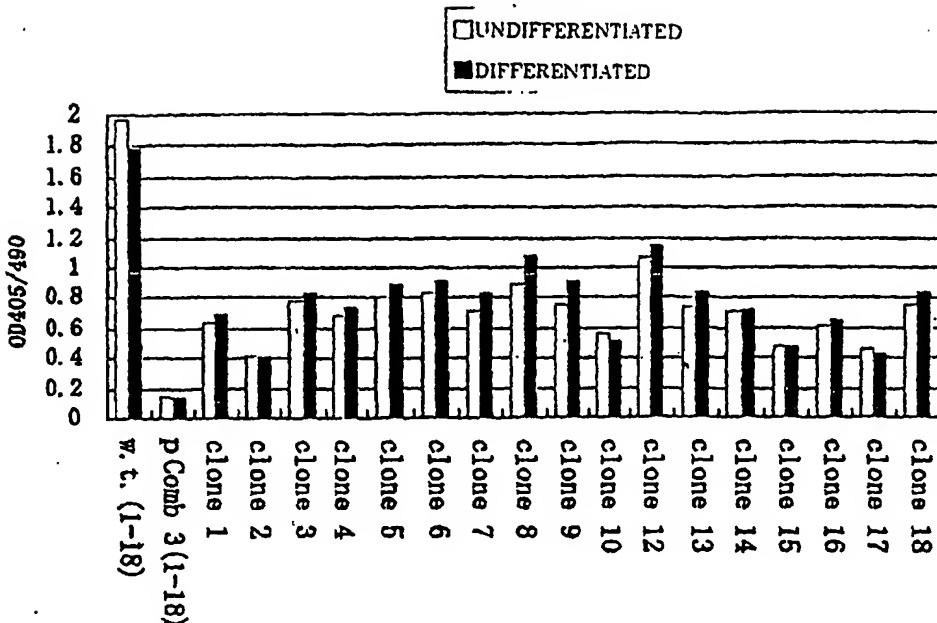


FIG. 10

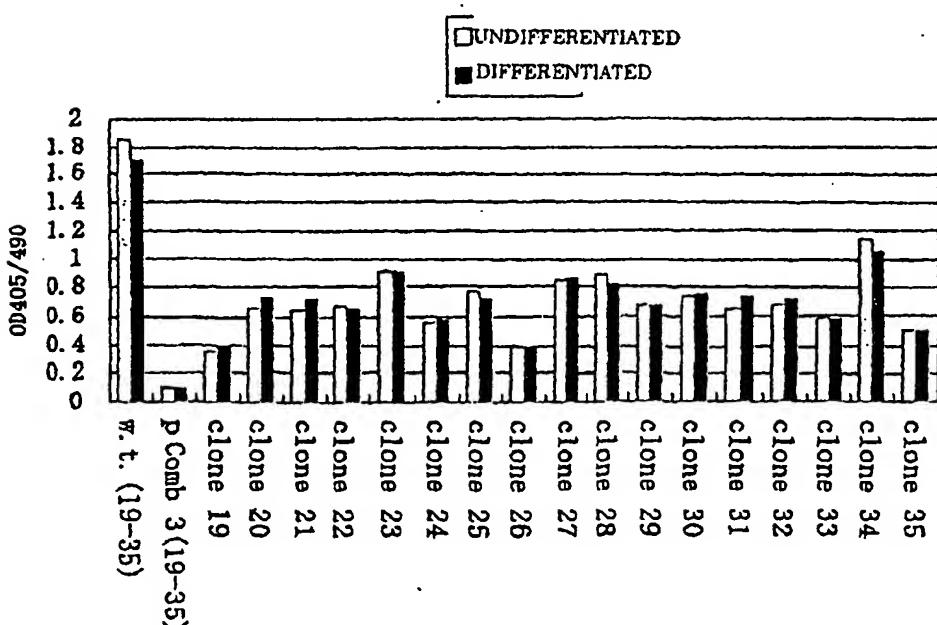
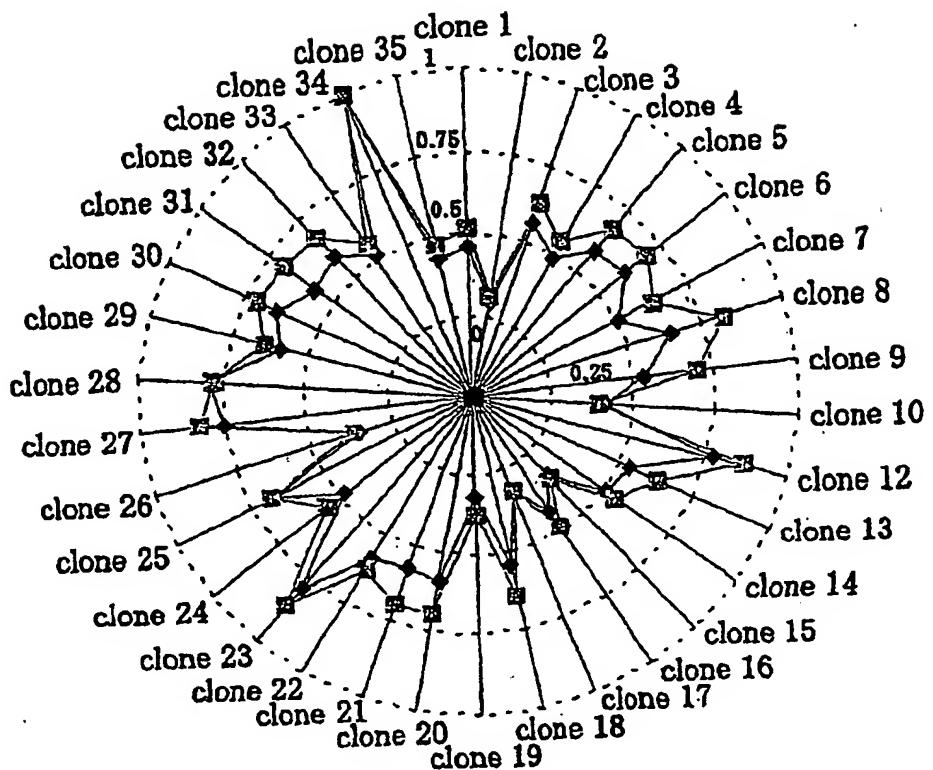


FIG. 11.



◆	UNDIFFERENTIATED
■	DIFFERENTIATED

FIG. 12

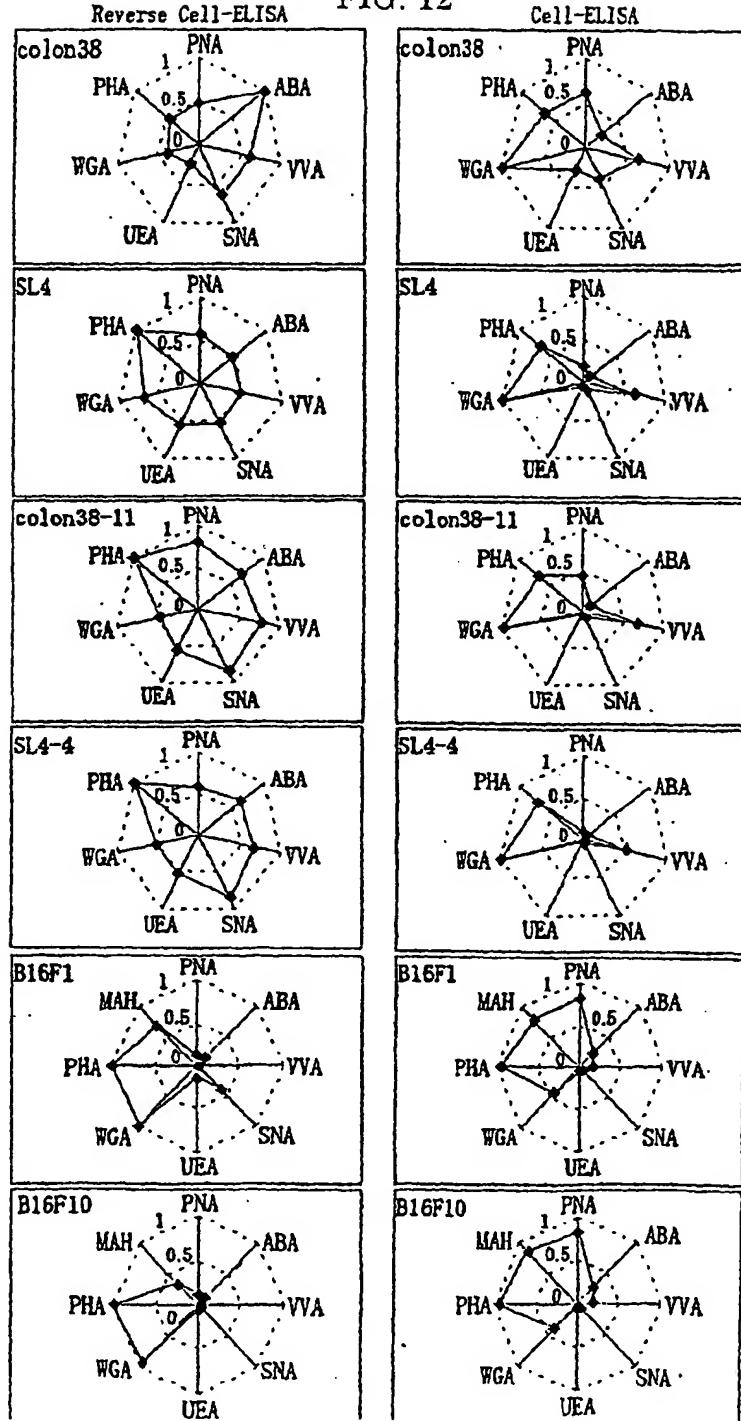


FIG. 13

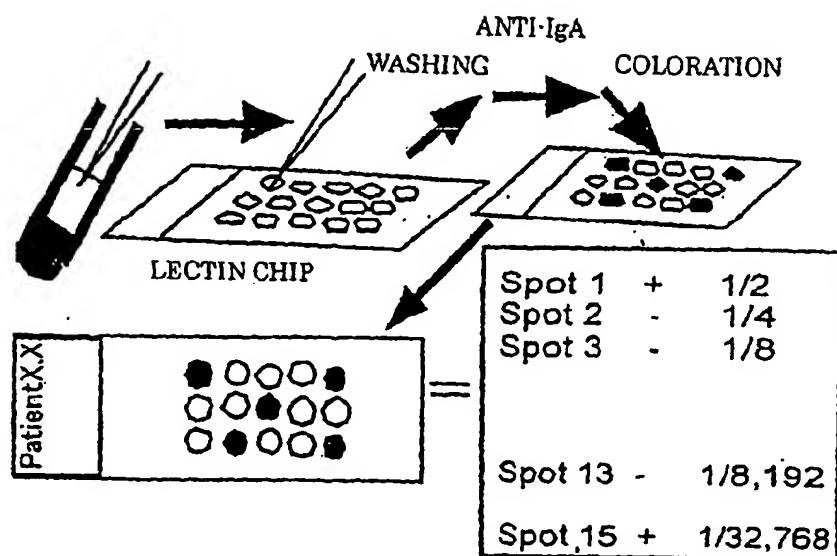
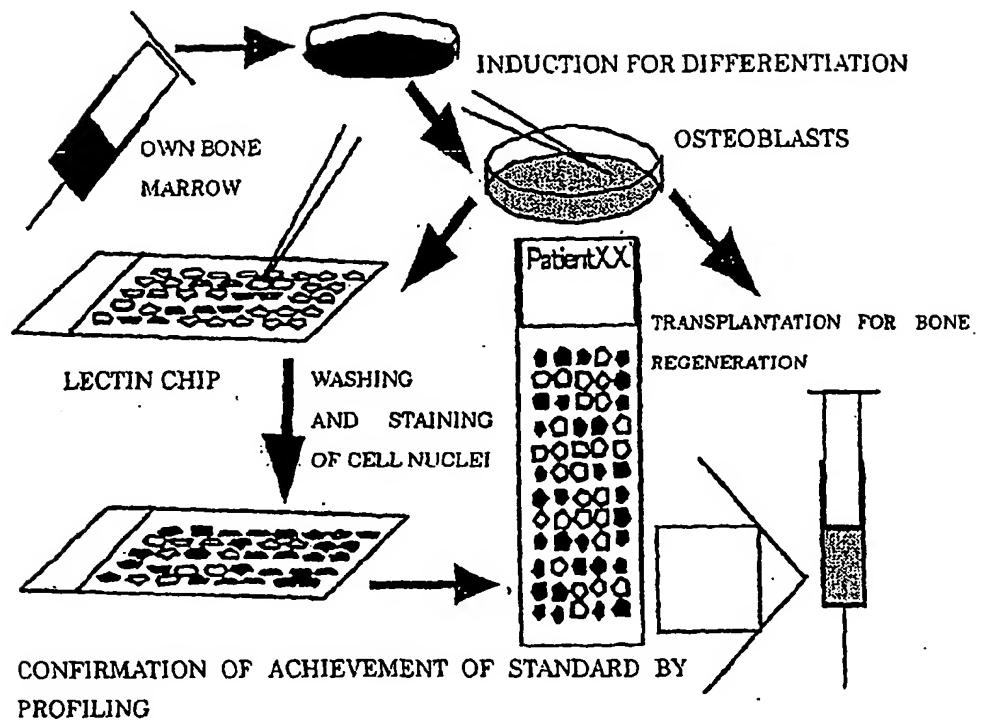


FIG. 14



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP02/01503

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl' C12N15/09, C12P21/02, C12Q1/02, G01N33/48

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl' C12N15/09, C12P21/02, C12Q1/02, G01N33/48

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
JICST FILE (JOIS), WPI (DIALOG), BIOSIS (DIALOG)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Mi-jung YIM et al., Generation of Libraries of Genetically Engineered Plant Lectins With Diverse Carbohydrate Specificities. 5th Annual Conference Of the Society for Glycobiology 2000, Vol.10, No.10, Pages 1094 to 1095	1
Y	Kazuo YAMAMOTO et al., Cyborg Lectins: Novel Leguminous Lectins with Unique Specificities. J.Biochem. 2000, Vol.127, pages 137 to 142	7,10 11,12
Y	Anne IMBERTY et al., An Unusual Carbohydrate Binding Site Revealed by the Structures of Two Maackia Murensis Lectins Complexed with Sialic Acid-containing Oligosaccharides. J.Biol.Chem. June 2000, Vol.275, No.23, pages 17541 to 17548	7,10

 Further documents are listed in the continuation of Box C.  See patent family annex.

* Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier document but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 12 April, 2002 (12.04.02)	Date of mailing of the international search report 23 April, 2002 (23.04.02)
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP02/01503

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Kazuo YAMAMOTO et al., Sialic acid-binding Motif of Maackia amurensis Lectins. J.Biochem. April, 1997, Vol.121, No.4, pages 756 to 761	7,10
P,X	Mijung YIM et al., Mutated Plant Lectin Library Useful to Identify Different Cells. Proc.Natl.Acad.Sci.USA February 2001, Vol.98, No.5, pages 2222 to 2225	1-12

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**